Ecotype Allelic Variation in C-to-U Editing Extent of a Mitochondrial Transcript Identifies RNA-Editing Quantitative Trait Loci in Arabidopsis[1]

Stéphane Bentolila[2], Anne-Laure Chateigner-Boutin[2,3], and Maureen R. Hanson*

Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853

In higher plants, RNA editing is a posttranscriptional process that converts C to U in organelle mRNAs. Although RNA editing in mitochondria occurs much more frequently than in chloroplasts, editing of exogenously supplied RNA substrates in vitro and in organello has shown that editing in the two organelles shares some common features. In particular, the 20 nucleotides upstream of the editing site play an important role in specifying the C to be edited. Biochemical approaches have allowed the identification of features of cis-sequences necessary for RNA editing to occur, but have failed to identify any of the components of the mitochondrial editing machinery. In order to implement a genetic approach for identification of editing factors, we have identified a polymorphism in the editing efficiency of a mitochondrial site between two ecotypes of Arabidopsis (*Arabidopsis thaliana*), Columbia (Col) and Landsberg *erecta* (L*er*). In rosette leaves, an editing site within the *ccb206* mitochondrial gene is more highly edited in Col than in L*er*. Depending on the development stage and tissue analyzed, the difference in editing extent varies between the two ecotypes; for example, in floral buds, editing extent does not differ. Single-point regression analysis of the editing efficiency in a sample of recombinant inbred lines derived from a cross between Col and L*er* allowed the identification of two quantitative trait loci controlling this trait. A member of the pentatricopeptide repeat protein family that carries a putative mitochondrial transit sequence has been identified near a major quantitative trait locus on chromosome 4.

RNA editing is a process that alters the genetic information at specific sites on RNA molecules. Editing has been described in a wide range of organisms from viruses to animals and plants. Several systems involving unrelated mechanisms seemed to have arisen separately during evolution (for review, see Gott and Emeson, 2000). In vascular plants, organelle transcripts are modified by C-to-U editing. The chloroplast genome of higher plants contains about 30 editing sites (Tsudzuki et al., 2001). In mitochondria, editing occurs more frequently; 427, 456, and 491 sites have been reported, respectively, in rapeseed (*Brassica napus*), Arabidopsis (*Arabidopsis thaliana*), and rice (*Oryza sativa*; Giege and Brennicke, 1999; Notsu et al., 2002; Handa, 2003).

Progress has recently been made in the characterization of chloroplast editing cis- and trans-factors since both in vivo (Chaudhuri and Maliga, 1996) and in vitro systems (Hirose and Sugiura, 2001; Miyamoto et al., 2002; Hegeman et al., 2005) have become available. Sequences important in the selection of the editable C have been found 5′ of C targets of editing in chloroplasts. Furthermore, clusters of two to five editing sites that share similar 5′ sequences have been identified. Evidence for functional significance of these clusters comes from their coinhibition in editing efficiency as a consequence of overexpression of one cluster member (Chateigner-Boutin and Hanson, 2002) as well as their developmental coregulation (Chateigner-Boutin and Hanson, 2003). A current hypothesis is that the editing sites within the same cluster share a nuclear-encoded protein recognition factor that may be expressed in limited quantities in some tissues at particular times in development, thus affecting editing efficiency.

In mitochondria, the lack of transformation methods and the absence of a reliable in vitro or in organello system, until recently, as well as the large number of editing sites, have slowed progress in identifying components of the molecular apparatus responsible for RNA editing. Electroporation of isolated wheat mitochondria with RNA substrates has revealed that, in mitochondria as in chloroplasts, 5′ sequences are important for specification of editing (Farré et al., 2001; Choury et al., 2004). The cis-requirements for the editing of a site in *atp9* mRNA have also been determined by editing within a pea mitochondrial extract (Takenaka and Brennicke, 2003; Takenaka et al., 2004). Results from these studies confirm the prominent role of the 20 nucleotides upstream of the editing site. Mitochondrion and plastid editing systems share

[1] This work was supported by grants from the National Institutes of Health (R01 GM50723) and the National Science Foundation (MCB 0344007).

[2] These authors contributed equally to the paper.

[3] Present address: Unité Mixte de Recherche 990, Institut National de la Recherche Agronomique (INRA)/Institut National Polytechnique-Ecole Nationale Supérieure Agronomique, "Genomique et Biotechnologie des Fruits," Centre INRA Toulouse Auzeville, Chemin de Borde Rouge, 31320 Castanet-Tolosan, France.

* Corresponding author; e-mail mrh5@cornell.edu; fax 607–255–6249.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Stéphane Bentolila (sb46@cornell.edu).

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.105.069013.

some other similarities. The 5′ portion of several mitochondrial site sequences exhibit similarities to putative cis-elements found in chloroplasts (Chateigner-Boutin and Hanson, 2002). This finding raised the possibility of shared recognition factors encoded by the nucleus and dually targeted to both organelles. However, it is unlikely that all factors are shared between the organelles, as it has been found that the editing sites within a portion of the *cox2* mitochondrial gene transcribed in transgenic chloroplasts were not edited by the chloroplast machinery (Sutton et al., 1995). Furthermore, presumably many more editing factors are needed in mitochondria than in chloroplasts.

Recently, an Arabidopsis gene that affects the editing efficiency of a chloroplast editing site has been identified (Kotera et al., 2005). While studying mutants affected in NADH dehydrogenase activity, the authors found that one such mutated allele, *ccr4*, also reduced editing of *ndhD*. The authors performed map-based cloning to discover that *CRR4* is a member of the pentatricopeptide repeat (PPR) gene family (Small and Peeters, 2000). *CRR4* is closely related in structure to *CRR2*, another PPR-containing gene involved in the intergenic RNA cleavage between *rps7* and *ndhB* (Hashimoto et al., 2003). The only structural difference between the proteins encoded by these two genes resides in a C-terminal 15-amino acid motif that is less conserved in CRR2. This 15-amino acid motif is too small to carry the catalytic activity of cytidine deamination that is necessary for the RNA-editing process. The authors therefore propose that CRR4 is a site recognition factor that interacts with the target RNA to recruit an editing enzyme, analogous to the interaction of the RNA-binding protein apobec-1 complementation factor (ACF) with a cytidine deaminase, APOBEC1, to the mammalian *apoB* editing site (Keegan et al., 2001).

A possible approach to identify editing factors is the genetic mapping of an editing polymorphism between two genotypes, followed by map-based cloning of the gene responsible for the polymorphism. Here, we present the identification of an editing polymorphism between two ecotypes of Arabidopsis, Columbia (Col) and Landsberg *erecta* (L*er*). The polymorphism occurs at a particular site within the mitochondrial transcript *ccb206* (*orf206*), which encodes a protein similar to a component of a heme transporter involved in cytochrome *c* biogenesis in photosynthetic bacteria (Schuster, 1994). This site is differentially edited between the two ecotypes, depending on the developmental stage and particular tissue. Recombinant inbred lines (RILs) generated from a cross between these two ecotypes allowed us to locate two quantitative trait loci (QTLs) linked to the editing efficiency of this site.

RESULTS

Editing Sites within *ccb206* Transcripts in Mitochondria

In order to find an editing site exhibiting a differential editing efficiency between the two ecotypes of Arabidopsis, Col and L*er*, we examined a gene that exhibits numerous editing sites. According to Giege and Brennicke (1999), the gene *ccb206* (NP_085482) has the largest number of edited Cs in Arabidopsis mitochondria; its open reading frame of 621 nucleotides was reported to contain 39 sites.

We examined the editing extent of all sites within *ccb206* by bulk sequencing of reverse transcription (RT)-PCR products obtained from rosette leaves and compared the cDNAs to the genomic sequence. Three sites originally described in Giege and Brennicke's (1999) survey of *ccb206* editing, which was performed on cultured cells of Arabidopsis ecotype Col, were not edited in rosette leaves. Furthermore, we found that *ccb206* transcripts exhibited an editing site not previously reported in Arabidopsis but described in rapeseed (Itani and Handa, 1998; Tables I and II). In Col and L*er* rosette leaves, 37 Cs are modified to Us in *ccb206* transcripts.

Differential Editing of *ccb206* C24 in Arabidopsis Col and L*er*

In the raw sequence data, the site *ccb206* C24 attracted our attention; at this position, the C/T was equally represented in Col, but C was prominent in L*er*. *ccb206* C24 is located in the first position within the codon CUG, modified to UUG by editing. Because both codons encode a Leu residue (position 136), C24 is a silent site. We confirmed that this site is partially edited in rosette leaves of both ecotypes by poisoned primer extension (PPE; Fig. 1). We found that C24 is edited in L*er* at an average of 23% and in Col at 51%. This site is differentially edited in the two ecotypes despite identical genomic sequences of the *ccb206* gene and comparable levels of the *ccb206* transcript (data

Table I. *Editing sites in ccb206 transcripts of rosette leaves of Arabidopsis ecotypes Col and Ler*

Position is in nucleotides from the A of the initiation codon.

Sites	Position	Sites	Position
C1	16	C20	338
C2	28	C21	367
C3	71	C22	379
C4	80	C23	380
C5	128	C24	406
C6	137	C25	424
C7	148	C26	428
C8	149	C27	467
C9	154	C28	475
C10	159	C29	476
C11	160	C30	485
C12	164	C31	512
C13	172	C32	514
C14	179	C33	551
C15	181	C34	554
C16	193	C35	566
C17	194	C36	569
C18	286	C37	596
C19	304		

Table II. *Differences in the editing sites of ccb206 in Arabidopsis ecotype Col between this report and Giege and Brennicke's report (1999)*

Sites 1, 2, and 3 were not found in rosette leaves, while C37 has not previously been reported.

Sites	Position	Sites	Position
1	75	3	188
2	78	C37	596

not shown). This is in contrast to a polymorphism previously detected in editing of a site in petunia (*Petunia hybrida*) *nad3. nad3* transcripts were much less abundant in a genotype exhibiting less editing of a particular C than in a genotype where the editing extent was higher (Lu and Hanson, 1992). In the petunia example, *nad3* transcripts in plants exhibiting the less edited, less abundant phenotype may have been less stable, so that turnover occurred before editing could be completed.

Developmental Regulation of *ccb206* C24 Editing in Arabidopsis Col But Not in L*er*

We investigated the editing efficiency of this mitochondrial site in different tissues of both ecotypes, knowing that the editing extent of partially edited sites can vary between tissues in plastids (Peeters and Hanson, 2002). We found that, in L*er*, C24 was edited at 23% to 24% in all investigated tissues (Fig. 1). On the contrary, Col C24 editing efficiency varied depending on the tissue analyzed. The highest editing level was found in rosette leaves, where the site is edited at 51%, and the lowest in floral buds, where C24 is edited only at 23%. In floral buds, C24 is edited to the same extent in the two ecotypes. We decided to continue our study on rosette leaves, where the difference was the most prominent.

Dominance of the L*er* Phenotype

To further analyze the genetic basis of this polymorphism, we crossed the two ecotypes in both combinations: Col or L*er* as a female. In both cases, we obtained several F_1 hybrids. Their hybrid status was checked using cleaved amplified polymorphic sequence markers (Konieczny and Ausubel, 1993; data not shown). The editing extent of *ccb206* C24 was investigated in rosette leaves of all F_1 hybrids. Hybrids generated from both cross combinations exhibit a L*er* phenotype (Fig. 2), which is therefore dominant over the Col phenotype. These data also suggest that any genes controlling the phenotype are located in the nucleus, since no maternal effect was observed in the editing extent of C24 in F_1 hybrids.

C24 Editing Extent of RILs

To map loci involved in the editing polymorphism between Col and L*er*, we assayed the editing extent of a population of RILs generated from a cross between the two ecotypes, with L*er* as the female parent (Lister and Dean, 1993). We analyzed the editing efficiency of *ccb206* C24 in the 30 RILs selected as having the highest frequency of recombination over the five chromosomes.

The PPE experiments were repeated three times and the average editing efficiencies are represented in Figure 3. The variance within the RILs was estimated to be 2.48, while the variance between the RILs was estimated to be 77.09. A likelihood ratio statistic test showed that the variance between the RILs was highly significantly different from zero. N1974 is the only outlier found in the RIL population with an average editing percentage of 17% compared to the average editing percentage of 23% for L*er*, the low-edited parent (Fig. 3). Finding only one outlier in the RIL population suggests that no QTL with an opposite effect to the ones found in the parent lines (namely, reduced editing for L*er* versus increased editing for Col) was

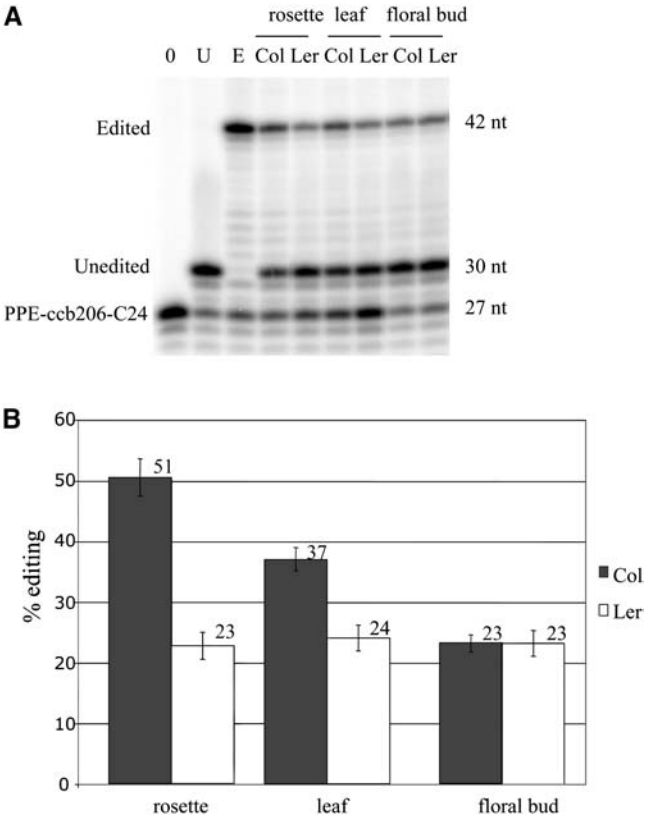

Figure 1. Editing extent of *ccb206* C24 in different tissues of Arabidopsis ecotype Col and L*er*. Editing efficiency was assayed by PPE. A, PPE was performed on an RT-PCR product from the radiolabeled oligo PPE-ccb206-C24; extension of the primer was poisoned by ddCTP incorporation. PPE products were resolved on a 12% sequencing gel, which was then exposed on a phosphorimager screen. B, The percentage of edited transcript was determined by quantifying the radioactivity associated with edited and unedited *ccb206* C24 using ImageQuant software (Molecular Dynamics). The editing efficiency is reported on the *y* axis and is the average of three independent PPE reactions. The error bars represent the SDS of the three measurements. 0, PPE without template indicating the size of the radiolabeled primer; U, PPE with a cloned (genomic) unedited PCR product; E, PPE with a cloned edited RT-PCR product.

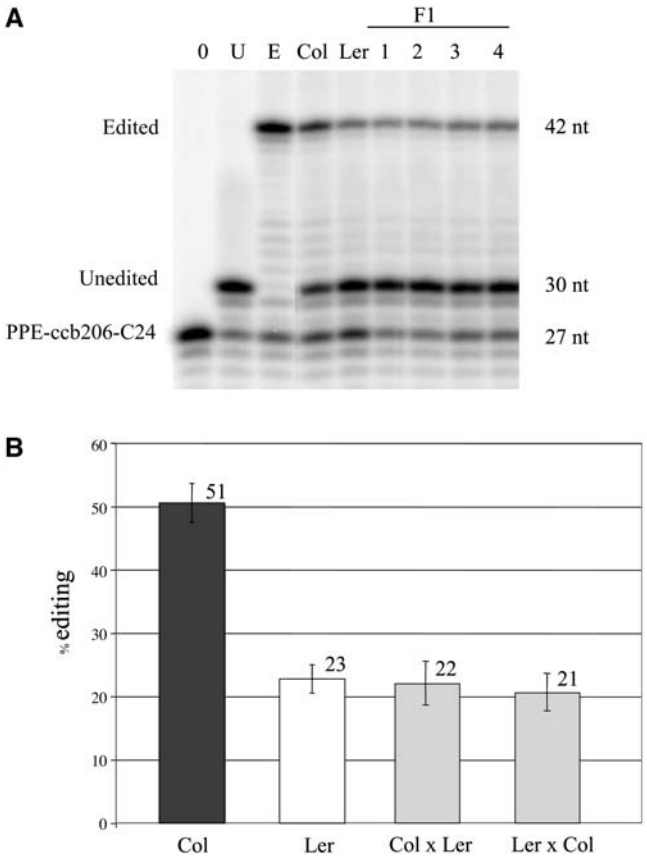

Figure 2. Editing extent of *ccb206* C24 in rosette leaves of F_1s generated from a cross between Arabidopsis ecotype Col and L*er*. A, PPE results obtained for the parents Col and L*er* and for four F_1 progeny. B, Editing efficiency of Col and L*er* and of F_1s generated from a cross between the two ecotypes in both combinations. The value for Col and L*er* is the same as the one reported in Figure 1B and comes from three different measurements. For the hybrids, the value is the average of assays of four different F_1 progeny for Col × L*er* and of three different F_1 progeny for L*er* × Col. The error bars represent the SDS of the measurements for each genotype. 0, PPE without template indicating the size of the radiolabeled primer; U, PPE with a cloned (genomic) unedited PCR product; E, PPE with a cloned edited RT-PCR product; Col × L*er*, F_1 generated from a cross with Col as a female; L*er* × Col, F_1 generated from a cross with L*er* as a female.

likely to be found. Some of the RILs exhibited a Col phenotype with an editing extent of C24 of about 50%, such as N1929 and N1901. Others have a L*er* phenotype, such as N1966, but most of the analyzed RILs exhibit an intermediate phenotype (Fig. 3). The fact that several RILs with editing efficiency comparable to the parental lines were found in a relatively small subsample of the RIL population argues for a relatively small number of QTLs controlling the editing efficiency of *ccb206* C24.

Localization of Two QTLs Linked to the Editing Efficiency of *ccb206* C24

The average editing efficiency of the RILs was used as a quantitative trait. In order to map QTLs with these RILs, we used 62 previously mapped markers. The markers span most of the genome at intervals <15 cM, except for a part of chromosome 3 (20.2 cM g4711-g4564-b) and chromosome 5 (27.2 cM g4715-b-m247; 18.3 cM g4028-m435). Identification of QTLs and estimation of their effects were obtained with three different analytical methods, single-point regression analysis (SPA), interval mapping (IM), and composite interval mapping (CIM). The three different approaches identified a major QTL on chromosome 4 colocalizing with the marker g6837 (Table III). The likelihood odds ratio (LOD) scores, which indicate how much more probable the data are to have arisen assuming the presence of a QTL than assuming its absence, are well above the empirical thresholds corresponding to an overall false positive rate of 1%. The amount of phenotypic variation explained by this QTL (R^2) ranges from 38% to 48%, depending on the method used to map it (Table III). At this peak marker, the average editing of C24 in Col-Col genotypes is 39%, whereas it is only 27% for L*er*-L*er* genotypes. Therefore, the Col allele is associated with an increase in C24 editing efficiency of about 6%, which represents the additive effect of this QTL. With the segregating population used in this study, it is not possible to estimate the dominance effect for any QTL, since all the RILs are homozygous for their whole genome.

The CIM alone identified another QTL with a smaller effect on chromosome 1 at the position occupied by the marker m532 (Table III). The amount of variation accounted for by this minor QTL amounts to 13%. As for the major QTL on chromosome 4, the Col allele is responsible for an increase in C24 editing efficiency of about 3% (Table III). The LOD of 3.39 associated with this QTL is above the overall 5% threshold of finding a false positive, but below the 1% threshold. The marker m532 also showed a peak for the SPA and the IM LODs above the nominal 5% significance (0.97 and 0.98, respectively), but well below the 5% experiment-wise threshold. Interestingly, a two-way marker interaction test between the marker g6837 (the major QTL) and every other marker in the dataset identified m532 (the minor QTL) as the second highest interaction ($F = 5.64$; $P < 0.025$). Even though the risk of type I error (declaring an interaction significant when it is not) is relatively high with such a low nominal threshold ($P = 0.79$), we think this result is worth mentioning.

Multiple regression analysis with the markers g6837 and m532 as factors showed that the proportion of phenotypic variation explained by the model amounts to 61%, which corresponds to the sum of the individual R^2 for each marker. From this result, we can conclude that the two QTLs are likely to have an independent action on the editing of *ccb206* C24.

Search for Mitochondrial Editing Site Homologs to *ccb206* C24 Controlled by the Same QTLs

Because editing of *ccb206* C24 is silent—editing does not change the encoded amino acid sequence—this site

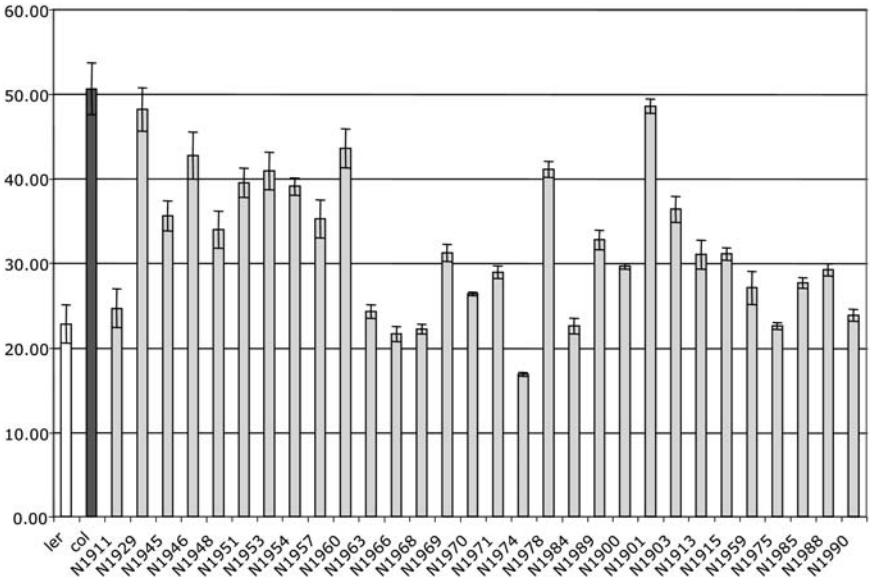

Figure 3. Editing extent of *ccb206* C24 in 30 RILs generated from a cross between Col and L*er*. The average of the editing percentage obtained from three different PPE experiments was determined as indicated in Figure 1. The error bars represent the SDs of the three measurements.

is likely to have no physiological importance. Why should silent editing occur? One possibility is that a recognition factor that operates on a nonsilent site happens to interact with C24 because of a fortuitous similarity in recognition cis-sequences between the two sites. The editing of C24 would then be a by-product resulting in editing that fails to alter the protein sequence. *ccb206* C24 could possibly cluster with more physiologically important sites, which would need editing to change the encoded amino acid residues. We attempted to find editing sites with a 5′ sequence similar to the one of *ccb206* C24 by using the 20 nucleotides upstream of the C target of editing in a search for short, nearly identical, sequences in the complete Arabidopsis mitochondrial editing database. Figure 4 shows a subset of the sites identified by Pratt software that we decided to investigate further. All these editing sites are nonsilent.

By investigating a site, we attempted to answer two questions. First, is the site differentially edited in Col and L*er* ecotypes? Second, are the QTLs controlling the editing of *ccb206* C24 also responsible for the observed difference in editing of the site being analyzed? Results are summarized in Table IV. All of the tested sites show a much higher editing extent than *ccb206* C24, with a majority of them being nearly fully edited. Only one site, *ccb203* (320), shows a significant difference in editing extent between the two parental genotypes. Like *ccb206* C24, the *ccb203* site is more edited in the Col ecotype than in the L*er* ecotype.

To find out whether the QTLs controlling *ccb206* C24 editing could also be involved in the editing of *ccb203* (320), we analyzed three RILs exhibiting the parental phenotypes for the editing of *ccb206* C24. N1901 and N1929, which possess the QTLs on chromosomes 1 and 4, are edited at about the same value as Col, while N1966, which lacks the QTLs, is edited at about the same value as L*er* (Table IV). If the same QTLs control any of the other sites, the expected result is that the editing of these particular RILs for this site should exhibit a parental phenotype. However, the observed values for *ccb203* (320) did not fit the expected values for any of the RILs. N1966, which exhibits low editing at the *ccb206* C24, exhibits a high *ccb203* (320) editing value close to Col, while N1901 and N1929, highly edited at the *ccb206* C24, are less edited at the *ccb203* (320) than N1966 (Table IV).

Differential Processing of the *ccb206* Transcript

Although we did not find a difference in the accumulation of the *ccb206* transcript in Col and L*er*, a

Table III. *QTLs controlling the editing efficiency of ccb206 C24 detected in a RIL population from a Col × Ler cross*

QTLs were detected by using three analytical methods: SPA, IM, and CIM. cM refers to the position of the marker on the chromosome (in cM). Source indicates the parental origin of the allele that increases the editing efficiency (Col). R^2 measures the proportion of variance explained by each marker. Effect is the additive effect contributed by each marker.

Chromosome	Marker[a]	cM	Source	*F*	SPA LOD[b]	R^2	Effect	LOD[b]	IM R^2	Effect	LOD[b]	CIM R^2	Effect
4	g6837	26.4	Col	26.85	**4.41**	0.48	6.15	**4.35**	0.45	5.95	**7.14**	0.38	5.61
1	m532	100.5	Col								3.39	0.13	3.33

[a]For chromosome 4, only the results for the peak marker are given. [b]Significance threshold levels ($P < 0.05$: SPA = 2.59; IM = 2.7; CIM = 2.7; $P < 0.01$: SPA = 3.35; IM = 4; CIM = 3.8) were determined by permutation tests (see "Materials and Methods" for more details). LODs greater than the $P < 0.01$ threshold are shown in bold.

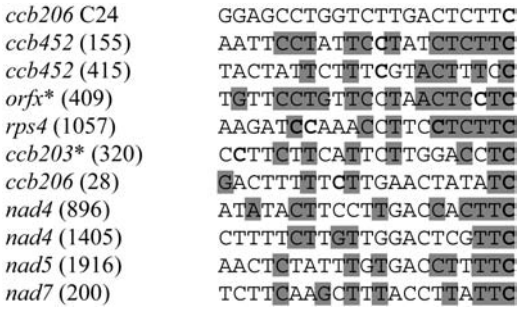

Figure 4. Editing sites sharing some similarity in their putative cis-recognition sequences with *ccb206* C24. These sites were identified by using Pratt software (see "Materials and Methods" for details). *ccb206* C24 is shown in the top line for reference. Identical nucleotides shared by other sites are shown in gray-shaded background. Cs targeted for editing are in bold. The number in parenthesis refers to the nucleotide position of the edited C in the coding sequence of each gene A from the initiation codon having the position 1. *, *orfx* and *ccb203* coding sequences are incomplete and do not start with an ATG codon.

different RNA-banding pattern was observed between the two ecotypes upon northern analysis (Fig. 5). It is thus possible that processing of the *ccb206* transcript specific to each ecotype is causing the observed difference in *ccb206* C24 editing. This hypothesis was tested by assessing the *ccb206* transcript pattern in eight RILs, four highly edited as Col and four poorly edited as L*er* (Fig. 5). At least three of these RILs show a *ccb206* RNA pattern contradicting this hypothesis. N1946, highly edited as Col, shows the same pattern as L*er*. N1974 and N1984, poorly edited as L*er*, show the same pattern as Col. (The pattern of N1929 is difficult to assess unequivocally on the northern presented in Fig. 5.) We can conclude from these data that processing of the *ccb206* transcript and editing of the *ccb206* C24 are independent.

DISCUSSION

We found a natural allelic variation in editing efficiency between two ecotypes of Arabidopsis. This editing polymorphism was found for a site in the mitochondrial transcript *ccb206*, the mRNA having the highest known number of editing sites in an Arabidopsis gene.

Editing of this particular site, *ccb206* C24, is partial in both ecotypes; however, in Col, but not in L*er*, the editing extent is differentially regulated in the tissues and developmental stages investigated. Crosses between the two ecotypes revealed that the L*er* phenotype is dominant over the Col phenotype and that the genes involved in this polymorphism are located in the nucleus.

One puzzling feature of the polymorphism we identified is that the less edited phenotype is dominant to the more edited phenotype. If the polymorphism reflects a difference in efficiency of recognition of a site or the ability to attract an editing enzyme to a site, then possibly the higher editing phenotype would be dominant. However, editing factors are present in limited quantities, as shown by the reduction in editing of endogenous genes when a plastid transgene carrying an editing site is overexpressed (Chateigner-Boutin and Hanson, 2002). Furthermore, members of the PPR gene family, some of which may encode editing factors, are known to exhibit transcripts with very low abundance (Lurin et al., 2004). It is possible that the dominant allele encodes a trans-factor that is more abundant than the factor encoded by the recessive allele. Furthermore, the presence of a trans-factor encoded by the dominant allele might result in low editing because it occupies the editing substrate, but does not bind as well to an editing enzyme that is required for completion of the editing reaction. According to this hypothesis, one or both QTL could represent factors that recognize editing sites despite the dominance of the less edited phenotype.

Alternatively, the QTLs could represent inhibitory factors that are directly or indirectly affecting editing efficiency. There is precedence for an inhibitory factor affecting C-to-U editing in the apoB gene, which is affected by the cytidine deaminase APOBEC-1 and ACF, an accessory factor. A regulatory factor termed GRY-RBP is an RNA-binding protein that shares 50% identity with ACF and inhibits both the binding of ACF

Table IV. *Editing extent of sites sharing similarity in their 20-nucleotide upstream sequences with ccb206 C24*

Editing was assayed on the parental genotypes Col and L*er* and on RILs showing a *ccb206 C24* editing phenotype very similar to the parents. ND, Not determined.

Site	Col	L*er*	N1901	N1929	N1966	N1974	N1946	N1975	N1984
ccb206 C24	51%	23%	49%	48%	21%	16%	43%	23%	23%
ccb452 (155)	93%	93%	93%	93%	93%	91%			
ccb452 (415)	94%	94%	96%	91%	92%	90%			
orfx (409)	87%	88%	84%	82%	86%	87%			
rps4 (1057)	77%	74%	76%	ND	73%	ND	77%	76%	68%
ccb203 (320)	82%	65%	76%	75%	80%	ND			
ccb206 (28)	94%	90%	95%	93%	93%	92%			
nad4 (896)	98%	98%	98%	98%	98%	98%			
nad4 (1,405)	93%	93%	93%	93%	92%	93%			
nad5 (1,916)	97%	98%	97%	98%	96%	97%			
nad7 (200)	97%	96%	97%	96%	96%	97%			

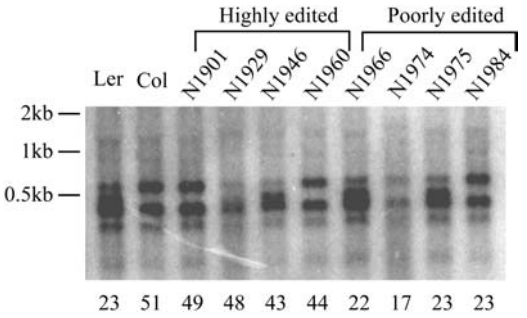

Figure 5. The *ccb206* transcript pattern does not cosegregate with *ccb206* C24 editing efficiency. Numbers below the figure refer to the editing extent of *ccb206* C24.

and the editing event (Blanc et al., 2001). In contrast to such a direct effect on editing, possibly the QTL might influence RNA-editing efficiency indirectly by modulating the abundance of a mitochondrial RNA species that competes for the *ccb206* C24 RNA. Perhaps a transcript carrying an editing site that clusters with the *ccb206* C24 sequence is more highly expressed in L*er* than in Col, leading to reduced editing because of a limited amount of the cluster's editing factors. The competing transcript's abundance could be regulated by one or both of the QTLs. A number of nuclear genes have been identified that control plant organelle transcript and protein abundance (for review, see Stern et al., 2004; Hanson and Bentolila, 2004).

We searched for a nonsilent editing site that could be the primary target of the QTLs identified in this study. Because *ccb206* C24 is a silent editing site that does not change the amino acid sequence encoded by the DNA, we speculated that there might be another nonsilent editing site sharing some sequence similarity in its upstream sequence with *ccb206* C24. Only one of the tested sites, *ccb203* (320), showed a significant difference between the Col and the L*er* ecotypes (Table IV). The editing assessment of some of the RILs for *ccb203* (320) did not support this polymorphism being controlled by the same QTLs as *ccb206* C24. Several reasons might explain why we have failed to identify a site that might be affected by the same QTLs as *ccb206* C24. We investigated only the most promising sites, based on visual inspection, among the sites identified by Pratt software. For instance, *orfx* (409), which has the closest match to *ccb206* C24, shares with the latter 13 identical nucleotides in its 20 upstream nucleotides (Fig. 4). Editing of the C in position −2 of *orfx* (409) makes its upstream sequence even more similar to *ccb206* C24. The same observation can be made for the sites *ccb452* (155) and (415), where editing of the C at

Table V. *Primers used in this study*

At-F and At-R were used to amplify RT-PCR products specific for each gene. PPE-gene-site primers were used to assay the editing extent of a specific site by PPE (Peeters and Hanson, 2002).

Gene	Primers
ccb206[a]	Atccb206F: ATGAGACGACTTTTTCTTGAACTATAT
	Atccb206R: TTAATCTTGTAAACTAATCGAGACC
	PPE-ccb206-C24: TTTTATTAGGGAGCCTGGTCTTGACTC
	PPE-ccb206-28: GATTGGTGTGGAGGAGAAGATCAG
ccb452	Atccb452F: GGTCCAACTACATAACTTTTTCTT
	Atccb452R: ATTATGAACTCCACGGAACTTTCT
	PPE-ccb452-155: TCGCTGACCTATCGCGTGCTAAAAAG
	PPE-ccb452-415: CCTTGAATGTAAATAGACCAAAAGAGAG
orfx[a]	AtOrfxF: CACTTTTAGCTTTGAATTACTTAT
	AtOrfxR: ATTGATAGTTACTTTGCCAGGTTC
	PPE-orfx-409: CCAAACATTGGGAACGACCCAGG
rps4[a]	Atrps4F: TCCCCATTAAGATTTCAAACTTGTCGTC
	Atrps4R: TTATATGTTTTGGCCACGTCCGTTTCTG
	PPE-rps4-1057: Atrps4-R
ccb203[a]	Atccb203F: TGGACACGGGGAGGGAGCAG
	Atccb203R: CATAACATAACGGGGCGGGGTTGC
	PPE-ccb203-320: CAGCATGGAAAAGTCACAATATTAAG
nad4	Atnad4F: AAGTGGTCTTATTCTGTGTCC
	Atnad4R: TTTGCCATGTTGCACTAAGTTACT
	PPE-nad4-896: CAATGATCTTCTTTAGATCGATCTG
	PPE-nad4-1405: GCAGTCCGGGAACACTTTGGGGTG
nad5	Atnad5F: GAAGGAAGCGCTATAATGACCACT
	Atnad5R: CGATCGATTATCTACCCAAGAAGA
	PPE-nad5-1916: GCAATGTTACTTGGTTCAACTCTATTTGTGACC
nad7	Atnad7F: ACTAGGAAAAGGCAAATCAAAAAT
	Atnad7R: ATCCACCTCTCCAAACACAATA
	PPE-nad7-200: GAGTACAAAACTTATCTTCAAGCTTTACC

[a]Transcripts of these genes do not contain introns; a control without reverse transcriptase was performed to check for DNA contamination.

position −9 makes the upstream sequences of these sites more similar to *ccb206* C24. Inversely, editing can decrease the similarity of an upstream sequence with the one from *ccb206* C24 as seen for *rps4* (1,057; edited Cs at positions −5 and −15). Some upstream sequences of the sites we investigated may appear to be poorly related to *ccb206* C24 because we have presented them without gaps for the purpose of simplicity (Fig. 4). But gaps were allowed in our screen with Pratt software. For instance, the introduction of a gap in the upstream sequence of *ccb206* C24 between the T at position −4 and the C at position −5 restores a contiguous stretch of six identical nucleotides (CTTGAC) between *ccb206* C24 and *nad4* (896). Gaps occur within the alignments of the upstream regions of known chloroplast editing-site clusters (Chateigner-Boutin and Hanson, 2002, 2003). The possible existence of gaps makes it difficult to select all possible candidate sites that might be related to *ccb206* C24.

Another explanation for the failure to find another editing site functionally related to *ccb206* C24 is our choice to screen the published editing database comprising 456 editing sites (Giege and Brennicke, 1999). As acknowledged by the authors, this database is not complete. About 31 kb of open reading frames >100 amino acids and 17 kb of introns have not been analyzed. Moreover, mitochondrial RNA was extracted from a cell-suspension culture in the study by Giege and Brennicke (1999). Editing is dependent upon the tissue analyzed and the environmental conditions in chloroplasts (Karcher and Bock, 1998; Nakajima and Mulligan, 2001; Peeters and Hanson, 2002) and in mitochondria (this study; Kurihara-Yonemoto and Handa, 2001). Some sites might thus have been overlooked because they are not edited in suspension cells. Although it is difficult to evaluate the number of such sites, it might be rather high. In an independent study covering 150 reported sites in seven mitochondrial genes, we found 17 new, fully edited sites and 22 new, partially edited sites (S. Bentolila, unpublished data).

A BLAST search for nearly identical short sequences to the 20 nucleotides upstream of the *ccb206* C24 in the Arabidopsis mitochondrial genome did not reveal any additional promising sequences to investigate further (data not shown). As discussed earlier, because editing can itself alter the RNA substrate that is targeted for further editing, possibly an RNA sequence highly similar to *ccb206* C24 could exist but not be recognizable by searching for matches to unedited genomic DNA.

Having failed to uncover a functional nonsilent site related to *ccb206* C24 does not invalidate our hypothesis about the existence of such a site. One possible reason we have not detected a polymorphism in a nonsilent site is that both of the proteins encoded by the QTL recognize the nonsilent site with strong affinity. If so, then high editing of the nonsilent site would occur in both ecotypes. If an accidental recognition of *ccb206* C24 is occurring, then the lower affinity of both factors for the silent site might have exposed differential binding of two factors encoded by different alleles, resulting in a differential editing. Thus, one of the sites identified by Pratt software and shown in Table IV might actually be the real target of the major QTL.

Could there be a selective advantage to editing a silent site? One possible scenario relates to the codon usage in the Arabidopsis mitochondrion, which shows a nearly systematic preference for codons ending with U for a specific amino acid (Giege and Brennicke, 1999). It has been speculated that this bias could partly explain the persistence of silent editing affecting the third position. However, while this theory might be valid as a factor in the persistence of silent editing at the species level, it is probably not relevant to the polymorphism we observed between the Col and L*er* ecotypes. It is unlikely that the overall codon usage in the mitochondrial genome of L*er* differs from codon usage in Col.

Recently, *CRR4*, a gene essential for the editing of a site that creates an initiation codon in the chloroplast gene *ndhD*, has been cloned (Kotera et al., 2005). *CRR4* is a member of the PPR family, one of the largest plant gene families first described in Arabidopsis (Small and Peeters, 2000). Although PPR-containing genes are found in all eukaryotes analyzed, they have greatly expanded in plants. According to the nature of the motifs found in the PPRs, this family can be divided into two subfamilies, the P subfamily and the PLS subfamily, which is specific to plants (Lurin et al., 2004). The PLS subfamily can be further divided into four subgroups, PLS, E, E+, and DYW, depending on the C-terminal domain found after the PPR motifs. Several lines of evidence point to an involvement of the PPR proteins in the control of gene expression in organelles. Most of the PPR proteins are predicted to be targeted to either mitochondria or plastids (Lurin et al., 2004). Positional cloning of cytoplasmic male sterility restorer genes, *Rf* in petunia (Bentolila et al., 2002), *Rfk* and *Rfo* in radish (Brown et al., 2003; Desloire et al., 2003; Koizuka et al., 2003), and *Rf-1* in rice (Komori et al., 2004) has shown that they all encode PPR proteins. Lurin et al. (2004) speculated that the PLS subfamily of PPR proteins, specific to plants, might be involved in RNA editing because this posttranscriptional process is unique to land plants. Two hundred PPR proteins belonging to the PLS subfamily have been identified in the Arabidopsis genome (Lurin et al., 2004). It is thus tempting to postulate that the QTLs reported in this study might indeed correspond to one of these PLS proteins. We looked for the presence of such a protein in the vicinity of g6837 on chromosome 4, the marker that colocalizes with the major QTL. We found one or two candidate genes for the major QTL depending on the length of the support interval obtained with the CIM, the method that gave the sharpest LOD peak of the three QTL mapping methods used in this study. The one LOD support interval for CIM, CIM1, covers 301 kb extending from the Arabidopsis Genome Initiative (AGI) map position 7,754 kb to the AGI map position 8,055 kb. In this interval, 67 predicted genes (http://www.ncbi.nlm.

nih.gov/mapview/maps.cgi?TAXID=3702&MAPS=default&CHR=IV) are found. Only one PPR-containing gene, *At4g13650*, is found in CIM1. At4g13650, the PPR protein encoded by this gene, is predicted to be targeted to mitochondria by Predotar (Small et al., 2004) and TargetP (Emanuelsson et al., 2000), two prediction programs designed to determine the subcellular localization of proteins. The two LOD support interval for CIM, CIM2, covers 553 kb extending from the AGI map position 7,554 kb to the AGI map position 8,107 kb. CIM2 comprises 141 predicted genes, of which two are PPR-containing genes. In addition to *At4g13650*, CIM2 also includes *At4g14050*. At4g14050, like At4g13650, is predicted to be targeted to mitochondria by both prediction programs. Both of these PPR proteins belong to the DYW subclass.

Two arguments support these PPR-containing genes, *At4g13650* and *At4g14050*, as promising candidates for the major QTL. First, their localization is predicted to be in the mitochondrion, as expected of a factor controlling the editing of a mitochondrial site. Second, both genes belong to the DYW subclass, which has recently been observed to exhibit a phylogenetic correlation with RNA editing in plants. The DYW domain was found in the Jungermanniales, the largest order of liverworts, where RNA editing is present. By contrast, the DYW domain was not found in the Marchantiales, where editing is absent (I. Small, personal communication).

CONCLUSION

In this initial study, we have mapped two QTLs linked to the difference in editing efficiency between two ecotypes of Arabidopsis. One of these QTLs lies on chromosome 4 and plays a major role, while the other QTL lies on chromosome 1 and shows a smaller effect. Our results prove that it is possible to map major editing QTLs even with a small population. Given the level of precision attached to QTL mapping, we must undertake finer mapping in order to clone the actual genes involved in this polymorphism. We intend also to utilize a candidate gene approach by analyzing the genes encoding the PPR proteins that are located in the vicinity of the QTLs. In this group, genes that encode PPR proteins predicted to be targeted to mitochondria and belonging to the PLS subfamily, specific to plants, are the most promising candidates. Mapping editing efficiency polymorphisms may facilitate the identification of the factors controlling RNA editing in plant mitochondria, leading to a greater understanding of the mechanism of action of genes that regulate this intriguing posttranscriptional process.

MATERIALS AND METHODS

Plant Material

Seeds of Arabidopsis (*Arabidopsis thaliana*) ecotype Col-4 (N933), L*er*-0 (NW20), and the RILs were obtained from the Nottingham Arabidopsis Stock Centre (NASC; http://nasc.nott.ac.uk). Seeds were sown in Metromix soil and transferred in the dark at 4°C for 5 d before growing in a growth chamber at 22°C in an 8-h-dark/16-h-light cycle.

Molecular Biology Techniques

Total RNA was extracted using a RNeasy Plant mini kit (Qiagen) and treated with a DNA-free kit (Ambion). First-strand cDNA was synthesized from 1.5 μg of DNA-free RNA for 1 h at 37°C with an Omniscript kit (Qiagen) using the gene-specific reverse primer At-R (Table V) following the manufacturer's protocol. Reactions without reverse transcriptase were performed to check for genomic DNA contamination whenever the transcript did not contain an intron. cDNA samples were amplified by a standard protocol (5 min at 94°C followed by 35 cycles of 94°C, 30 s, 50°C, 30 s, 72°C, 2 min) in a PTC-200 thermal cycler (MJ Research) using gene-specific forward and reverse primers At-F and At-R (Table V).

Northern-blot analysis (Sambrook and Russell, 2001) was conducted with 10 μg of total RNA, using a radiolabeled *ccb206* RT-PCR product as a probe.

Genomic DNA was extracted (Dellaporta et al., 1983) and cleaved amplified polymorphic sequence analysis of F_1 hybrids was performed (Konieczny and Ausubel, 1993) using alcohol dehydrogenase primers.

PPE of RT-PCR products and determination of editing efficiency were conducted as previously described (Peeters and Hanson, 2002) using the primers PPE-gene-site (Table V). To confirm the results, all experiments were performed three times.

Statistical Analysis

The estimation of the variance of the editing efficiency between the RILs and within the RILs was done by fitting a variance component model using the mixed procedure in the SPSS software version 13.0 for PC. A likelihood ratio statistic test performed by using SPSS allowed the determination as to whether the variance between the RILs was significantly different from zero.

The chromosomal location of QTLs for editing efficiency was determined by three analytical methods, SPA using the Qgene computer program version 3.07 (Nelson, 1997), IM, and CIM, using Windows QTL Cartographer version 2.5 (Wang et al. 2005). CIM was implemented by using the standard model 6 (as recommended by the authors of the software as default) specifying five cofactors to control for genetic background and a window size of 10 cM that blocked out a region of the genome on either side of the markers flanking the test site. Since the flanking regions are tightly linked to the test site, they have to be eliminated from the background markers in order to be able to detect a signal from the test site itself. The specific cofactors used in the model 6 were obtained by two different methods, a forward regression and a forward-backward stepwise regression with $P_{in} = 0.01$ and $P_{out} = 0.01$. Because the QTLs identified with these two methods are the same, only the data obtained with the forward regression will be presented. The walking speed for both CIM and IM was set up to 2 cM.

We used permutation tests to establish experiment-wise significance thresholds for the three different analyses, as suggested by Churchill and Doerge (1994). For SPA, 10,000 permutations allowed the determination of a threshold of $F > 13.66$ or LOD > 2.59 that corresponds to an experiment-wise threshold of $P < 0.05$. An experiment-wise threshold of $P < 0.05$ was also chosen for the IM and the CIM and corresponds to LODs > 2.7 for both analyses (1,000 permutations). A more stringent experiment-wise threshold of $P < 0.01$ was computed in the same way for the three analyses and corresponds to LOD > 3.35 (SPA), LOD > 4 (IM), and LOD > 3.8 (CIM).

The proportion of observed phenotypic variance attributable to a particular QTL was estimated by the coefficient of determination (R^2) from the corresponding linear model (SPA) and using maximum likelihood for IM and CIM.

Pairwise epistatic interactions between QTLs were tested by a two-way ANOVA using Qgene software. Multiple regression analysis was also conducted by using Qgene software.

The search for mitochondrial editing sites homologous to *ccb206* C24 was performed by using the Pratt computer program (Jonassen, 1997). Pratt software was developed to identify a pattern in a DNA sequence. The pattern may contain a flexible number of gaps and ambiguous symbols at certain positions. The user supplies a set of unaligned sequences and the minimum number of sequences to match the pattern. During the first phase, patterns are constrained to the pattern class defined by the set of options given by the user. The pattern graph is constructed using a depth-first search algorithm. A search is done for the highest scoring patterns in the class that is derived from

the pattern graph. The highest scoring patterns found during this search are input to a heuristic pattern refinement algorithm, where more ambiguous pattern components can be used. In case flexible gaps are allowed, it is not guaranteed that the heuristic optimization algorithm will find the highest scoring pattern. The program analyzed 456 sequence segments of length 20 containing the 20 nucleotides upstream of each reported RNA-editing site (Giege and Brennicke, 1999). The input parameters that were used are as follows: CM = 5, C% = 1.1, pp = start, PL = 20, PN = 20, PX = 3, FN = 4, FL = 3, FP = 10, BI = off, BN = 20, S = info, G = query, GF = seq.fasta, E = 2, R = on, RG = off, OP = on, ON = 200, OA = 200, M = on, MR = on, MV = on. The input format can be found at http://us.expasy.org/tools/pratt/pratt-doc.html%23how1. The file seq.fasta contains a segment of the sequence *ccb206* C24 ggagcctggtcttgactcttc. Highly conserved patterns with high scores were chosen as candidates for further screening.

ACKNOWLEDGMENTS

We thank Dr. Tamara Galor from the Theory Center at Cornell University for performing the search for sites similar to *ccb206* C24 with the Pratt software. We wish to express our thanks to Susan McCouch and Clare Nelson for their helpful comments, and to Françoise Vermeylen from the office of Statistical Consulting at Cornell University for her help with the statistical analyses.

Received July 27, 2005; revised September 13, 2005; accepted September 23, 2005; published November 11, 2005.

LITERATURE CITED

Bentolila S, Alfonso AA, Hanson MR (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. Proc Natl Acad Sci USA **99:** 10887–10892

Blanc V, Navaratnam N, Henderson JO, Anant S, Kennedy S, Jarmuz A, Scott J, Davidson NO (2001) Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing. J Biol Chem **276:** 10272–10283

Brown GG, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang J, Cheung WY, Landry BS (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J **35:** 262–272

Chateigner-Boutin AL, Hanson MR (2002) Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol **22:** 8448–8456

Chateigner-Boutin AL, Hanson MR (2003) Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar cis-elements. Nucleic Acids Res **31:** 2586–2594

Chaudhuri S, Maliga P (1996) Sequences directing C to U editing of the plastid psbL mRNA are located within a 22 nucleotide segment spanning the editing site. EMBO J **15:** 5958–5964

Choury D, Farre JC, Jordana X, Araya A (2004) Different patterns in the recognition of editing sites in plant mitochondria. Nucleic Acids Res **32:** 6397–6406

Churchill G, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep **1:** 19

Desloire S, Gherbi H, Laloui W, Marhadour S, Clouet V, Cattolico L, Falentin C, Giancola S, Renard M, Budar F, et al (2003) Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatricopeptide-repeat protein family. EMBO Rep **4:** 588–594

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol **300:** 1005–1016

Farré JC, Leon G, Jordana X, Araya A (2001) Cis recognition elements in plant mitochondrion RNA editing. Mol Cell Biol **21:** 6731–6737

Giege P, Brennicke A (1999) RNA editing in Arabidopsis mitochondria affects 441 C to U changes in ORFs. Proc Natl Acad Sci USA **96:** 15324–15329

Gott JM, Emeson RB (2000) Functions and mechanisms of RNA editing. Annu Rev Genet **34:** 499–531

Handa H (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucleic Acids Res **31:** 5907–5916

Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell (Suppl) **16:** S154–S169

Hashimoto M, Endo T, Peltier G, Tasaka M, Shikanai T (2003) A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in Arabidopsis. Plant J **36:** 541–549

Hegeman CE, Hayes ML, Hanson MR (2005) Substrate and cofactor requirements for RNA editing of chloroplast transcripts in Arabidopsis *in vitro*. Plant J **42:** 124–132

Hirose T, Sugiura M (2001) Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. EMBO J **20:** 1144–1152

Itani K, Handa H (1998) Rapeseed mitochondrial *ccb206*, a gene involved in cytochrome c biogenesis, is co-transcribed with the *nad3* and *rps12* genes: organization, transcription, and RNA editing of the *nad3/rps12/ccb206* locus. Curr Genet **34:** 318–325

Jonassen I (1997) Efficient discovery of conserved patterns using a pattern graph. Comput Appl Biosci **13:** 509–522

Karcher D, Bock R (1998) Site-selective inhibition of plastid RNA editing by heat shock and antibiotics: a role for plastid translation in RNA editing. Nucleic Acids Res **26:** 1185–1190

Keegan LP, Gallo A, O'Connell MA (2001) The many roles of an RNA editor. Nat Rev Genet **2:** 869–878

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T, Kawasaki S, Imamura J (2003) Genetic characterization of a pentatricopeptide repeat protein gene, *orf687*, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J **34:** 407–415

Komori T, Ohta S, Murai N, Takakura Y, Kuraya Y, Suzuki S, Hiei Y, Imaseki H, Nitta N (2004) Map-based cloning of a fertility restorer gene, *Rf-1*, in rice (*Oryza sativa* L.). Plant J **37:** 315–325

Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J **4:** 403–410

Kotera E, Tasaka M, Shikanai T (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature **433:** 326–330

Kurihara-Yonemoto S, Handa H (2001) Low temperature affects the processing pattern and RNA editing status of the mitochondrial *cox2* transcripts in wheat. Curr Genet **40:** 203–208

Lister C, Dean C (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in Arabidopsis markers in *Arabidopsis thaliana*. Plant J **4:** 745–750

Lu B, Hanson MR (1992) A single nuclear gene specifies the abundance and extent of RNA editing of a plant mitochondrial transcript. Nucleic Acids Res **20:** 5699–5703

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, et al (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16:** 2089–2103

Miyamoto T, Obokata J, Sugiura M (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of *cis*-acting elements and *trans*-acting factors involved in RNA editing in tobacco and pea chloroplasts. Mol Cell Biol **22:** 6726–6734

Nakajima Y, Mulligan RM (2001) Heat stress results in incomplete C-to-U editing of maize chloroplast mRNAs and correlates with changes in chloroplast transcription rate. Curr Genet **40:** 209–213

Nelson JC (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed **3:** 239–245

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics **268:** 434–445

Peeters NM, Hanson MR (2002) Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA **8:** 497–511

Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schuster W (1994) The highly edited *orf206* in Oenothera mitochondria may encode a component of a heme transporter involved in cytochrome c biogenesis. Plant Mol Biol **25:** 33–42

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics **4:** 1581–1590

Small ID, Peeters N (2000) The PPR motif—a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci **25:** 46–47

Stern DB, Hanson MR, Barkan A (2004) Genetics and genomics of chloroplast biogenesis: maize as a model system. Trends Plant Sci **9:** 293–301

Sutton CA, Zoubenko OV, Hanson MR, Maliga P (1995) A plant mitochondrial sequence transcribed in transgenic tobacco chloroplasts is not edited. Mol Cell Biol **15:** 1377–1381

Takenaka M, Brennicke A (2003) In vitro RNA editing in pea mitochondria requires NTP or dNTP, suggesting involvement of an RNA helicase. J Biol Chem **278:** 47526–47533

Takenaka M, Neuwirt J, Brennicke A (2004) Complex *cis*-elements determine an RNA editing site in pea mitochondria. Nucleic Acids Res **32:** 4137–4144

Tsudzuki T, Wakasugi T, Sugiura M (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. J Mol Evol **53:** 327–332

Wang S, Basten CJ, Zeng ZB (2005) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC